

6 (261)

From: Borin, Michael
Sent: Tuesday, February 26, 2002 1:59 PM
To: STIC-Biotech/ChemLib
Subject: Search request: 09/540325

Examiner: M.Borin
CM1 12A01
AU: 1631; Mailbox 12D01

Tel.: 305-4506

235
RE: 09/540325, cyanidium nucl. acids

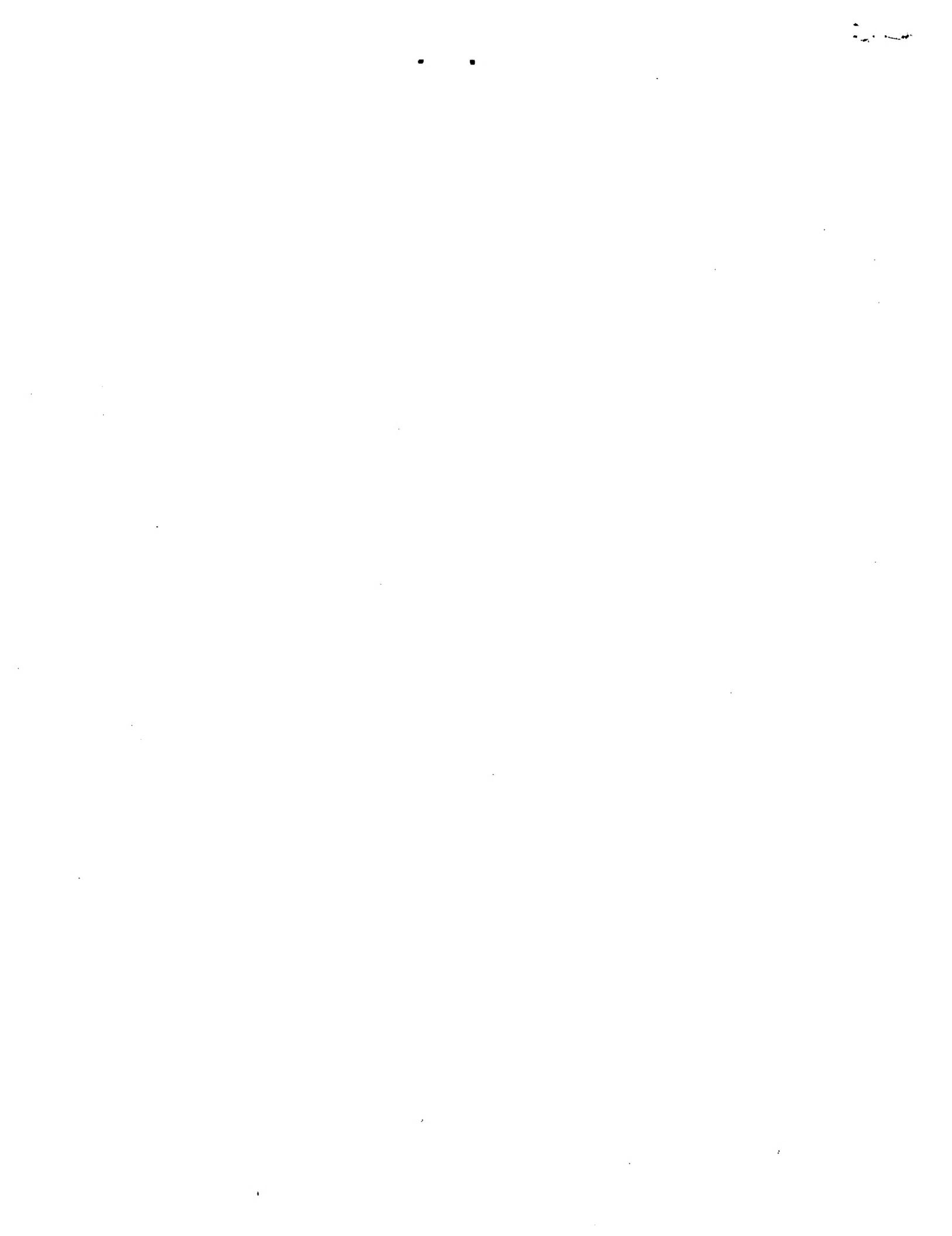
Please conduct search of polynucleotide SEQ ID 1 against the commercial and interference databases.

Thank you

Searcher: Paula Sheppard
Phone: 308-4499
Location: _____
Date Picked Up: 3/02/02
Date Completed: 3/04/04
Searcher Prep/Review: 10
Clerical: _____
Online time: 5

TYPE OF SEARCH:
NA Sequences: 1
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model1
Run on: March 2, 2002, 20:03:35 ; Search time 1441.26 Seconds
(without alignments)
4670.115 Million cell updates/sec

Title: US-09-540-235-1
Perfect score: 408

Sequence: 1 acggcgatggat... acatccatctaatgttta 408
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

* Database : GenBmlb:
1: gb_bp: *
2: gb_hg: *
3: gb_in: *
4: gb_on: *
5: gb_ov: *
6: gb_patt: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_cm: *
20: em_or: *
21: em_ov: *
22: em_patt: *
23: em_ph: *
24: em_pl: *
25: em_ro: *
26: em_sts: *
27: em_sy: *
28: em_un: *
29: em_vl: *
30: em_htgo_hum: *
31: em_htgo_inv: *
32: em_htgo_rod: *
33: em_htg_hum: *
34: em_htg_inv: *
35: em_htg_rod: *
36: em_htg_other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARES

Result No.	Score	Match Length	DB	ID	Query	Description
c 1	130	31.9	51860	8	AB013390	AB013390 Arabidops
c 2	124.2	30.4	729	8	AF370158	AF370158 Arabidops
c 3	124.2	30.4	9771	8	ATTF4	ATTF4
c 4	124.2	30.4	108158	8	ATTF4	ATTF4
c 5	117.8	28.9	11941	8	YSP72HB	YSP72HB
c 6	117.8	28.9	14147	8	SPBC29B5	SPBC29B5
c 7	110.2	27.0	97911	3	AC091511	AC091511 Leishmani
c 8	109.8	26.9	979	11	CNS061D1T	AL400167 T7 end of AX07328 Sequence
c 9	109.8	26.9	1348	5	CNS061D1T	AL113932 Borytis
c 10	106.6	26.1	480	8	CNS01B90	AL115461 Borytis
c 11	106.6	26.1	506	8	CNS01CFH	AL11303 Borytis
c 12	106.6	26.1	540	8	CNS0197Z	AL111947 Borytis
c 13	106.6	26.1	540	8	CNS019PV	AL11536 Borytis
c 14	106.6	26.1	540	8	CNS01AOA	AL11517 Borytis
c 15	106.6	26.1	540	8	CNS0194W	AL11386 Borytis
c 16	106.6	26.1	540	8	CNS01C0M	AL114653 Borytis
c 17	106.6	26.1	540	8	CNS01C4L	AL115790 Borytis
c 18	106.6	26.1	540	8	CNS01D4C	AL116041 Borytis
c 19	106.6	26.1	566	8	CNS01C0H1	AL111947 Borytis
c 20	106.4	26.1	2104	8	SCYGR034W	AL11386 Borytis
c 21	105	26.7	8	8	CNS019VN	AL112155 Borytis
c 22	102.2	25.0	626	8	D78495	DR8495 Brassica rapa
c 23	101.2	24.8	137129	2	HSAC001236	AC001236 Homo sapi
c 24	101.2	24.8	140977	9	AC002536	AC002536 Human Chr
c 25	99.6	24.4	66715	2	AC091012	AC091012 Homo sapi
c 26	99.6	24.4	160847	2	AC026645	AC026645 Homo sapi
c 27	96.2	508	10	10	RRPL26	X14671 Rat liver m
c 28	95.8	23.5	1331	6	AX073088	AX073088 Sequence
c 29	95.8	23.5	37541	8	YSC18300	U19028 Saccharomyces cerevisiae
c 30	95.8	23.4	10383	9	HUMRPL26X	U67497 Metathoracocarid L07287 Human ribs
c 31	94.4	23.1	2704	9	AC02431	AC02431 Homo sapi
c 32	92	22.5	417	8	AF03540	AF03540 Zea mays
c 33	91.4	22.4	141827	2	AC02913	AP02913 Oryza sativa
c 34	90.4	22.2	773	9	HSR26AA	X69392 H.sapiens
c 35	88.6	21.7	959	11	CNS061JUZ	AL40281 T7 end of
c 36	88.6	21.7	545	10	MM126MRN	X806595 M.musculus
c 37	87.2	21.4	173693	9	AC02431	AC02431 Homo sapi
c 38	87.2	21.4	183407	2	AC021572	AC021572 Homo sapi
c 39	87.2	21.4	183407	2	AC026478	AC026478 Mus musculus
c 40	87.2	21.4	190498	2	AC079871	AC079871 Mus musculus
c 41	87.2	21.4	209384	2	AC016618	AC016618 Homo sapi
c 42	85.6	21.4	147210	9	AC008391	AC008391 Homo sapi
c 43	84.8	20.8	172802	9	AC078817	AC078817 Homo sapi
c 44	84.8	20.8	182301	9	AC012498	AC012498 Homo sapi
c 45	84	20.6	110000	2	LMFICHR36_22	Continuation (23 o
ALIGMENTS						
RESULT	1					
AB013390/_C						
LOCUS	AB013390	51860 bp	DNA	PLN	27-DEC-2000	
DEFINITION	Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone:K919.					
ACCESSION	AB013390	BA00015				
VERSION	AB013390.1	GI:3128137				
KEYWORDS						
SOURCE	Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui TAC					
REFERENCE	Arabidopsis thaliana					
AUTHORS	Kotani,H., Nakamura,Y., Sato,S., Asamizu,E., Kaneko,T., Miyajima,N. and Tabata,S.					
TITLE	Structural analysis of Arabidopsis thaliana chromosome 5. VI. Sequence features of the regions of 11,367,185 bp covered by 19 physically assigned P1 and TAC clones					
JOURNAL	DNA Res. 5 (3), 203-216 (1998)					

FEATURES	COMMENT
source	Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3', 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/ . Location/Qualifiers
1. . 108158	
/variety="Arabidopsis thaliana"	
/db_xref="taxon:3702"	
/chromosome="3"	
gene	
7566. . 8213	
/gene="F3A4.10"	
/db_xref="8213	
/note="overlap to BAC T16K5, please refer to Accession EMBL:AL1329655 for analysis and annotation"	
7566. . 8213	
/gene="F3A4.10"	
/note="strong similarity to Cys2/His2-t-type zinc finger protein 1 - Arabidopsis thaliana, EMBL: AB030731; Contains zinc finger, C2H2 type, domain AA96-116; Zinc finger, C2H2 type, domain AA151-171"	
/codon_start=1	
/product="zinc-finger-like protein"	
/db_xref="GI:6522914"	
/translation="MALDILNSPNSITTTTAPPPLRCLEDETERENLESWIKRKTRK	
/transl_id="CAB62101.1"	
exon	
7566. . 8213	
/gene="F3A4.10"	
exon	
11785. . 11968	
/gene="F3A4.20"	
/number=1	
11785. . 12622	
/gene="F3A4.20"	
/join="11785. . 11968. 12063. . 12622)	
gene	
7566. . 8213	
/gene="F3A4.20"	
/note="similarity to hypothetical protein - Arabidopsis thaliana, EMBL:CAB38293"	
/codon_start=1	
/product="putative protein"	
/protein_id="CAB62102.1"	
/db_xref="GI:6522915"	
/translation="MSNCNSRVLRKGCSENCLRPCIOWLESPEAOGHATWFAKFFG	
RAGIMSFIAVPESCOPCAQQLQSLIEAACGRKVNPNGAVGLWGWNNVQQAETVIL	
RGGSKKPIPELNGGGFASRSPSPDEASISCTEMLNKRKADSDRNTYHCRSS	
RSRSSTASPPKKRSLQSQPSSELSLIPPIYKPLTFKEDPSMSYSEESVYTIV	
FONNNAGDRYVRCGGGGGAGTAKLNLFA"	
intron	
11969. . 12062	
/gene="F3A4.20"	
/number=1	
12063. . 12622	
/gene="F3A4.20"	
/number=2	
exon	
19890. . 21122	
/gene="F3A4.30"	
/number=1	
19890. . 21122	
/gene="F3A4.30"	
/number=1	
19890. . 21122	
/gene="F3A4.30"	
/note="similarity to lateral suppressor protein - Lycopersicon esculentum, EMBL:AF098674"	
gene	
CD5	
/db_xref="GI:6522916"	
/translation="MKTKLNLNTRPSPKPLRGCGDANPMEQLLHCATAIDNSDA	
LTHQILWVWNLLAPPGDSRDLTAFRLRSRAVSKTPILSSTSFFPDADELRF	
SWVLEAFLDLPWIRFGFIAANAAITAVGYSVTHIDLSLTHCQMOIPLIDAMAS	
RNLKPPPLKLITVSSSDFFPINTSYEGLVNFATRTRNITMEFTIVPSDGF	
FSSLQLQRLTPSSFNLALVNNMLHRYPEEPLTSSSRLRTVTLKLQKLSNPRI	
gene	
exon	
1. . 27649. . 28797	
/note="similarity to several hypothetical proteins -	
CDS	

TAPSRSPYVPLKALIGSSEBVNSNLVLTDEINTDGDIDPKFLISVRKFLMGSPNSXXV
XKHERRKLPGPRVYLXLRASVDEENKERLSPSDPVEVESVYVDNGELSNQKVERTV
EISKEQNDPAGSQYL"

CDS

complement (9759..10061)
<note>"unnamed protein product"
<protein_start=1>
<protein_id=BAA12201..1>
<db_xref="GI:1827311">
<translation>MARRTRPSPGRRFSCLTXXELGDPIKPNFLTEIKNIGSVPSPV
FVSKSVNTKILMSELPIAKSGTYDGEAGASVCNESCESSSLENPKSVALTPNL
10774..11364
<note>"unnamed protein product"
<codon_start=1>
<protein_id=BAA12202..1>
<db_xref="GI:1827512">
<translation>MISSLAKTREADXWSEEKTNYNELLSNPCKLCELDHFDSF
PHQODDLKNUYLAEBSASLFLDPLNULWQKQVFSKIGKPKLWMPISFHVQRFDP
LWQNTSTAEFTSYDRETTYDNTTILKXTAPSRTAQNWLFMHLIEIVSYLFPH
AHLDQYSLDDIYASVVERHPELGIVORKGFLQESRI

BASE COUNT ORIGIN
3774 a 2224 c 2065 g 3755 t 123 others

Query Match 28.9%; Score 117.8; DB 8; Length 11941;
Best Local Similarity 61.0%; Fred. No. 6.4e-21;

Matches 191; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

Query 96 atgaagtattccaaatgttatcgctgtctatggaaagcggaaatctttca 155
Db 4235 ATGAACTCTCGGAGTCACTACCTCTCGTCAAGCAGAAAGGTCTACCTGGC 4176

Query 156 gacccatctttgtacgacaaatctatcgatggacccctgtccaaaggactatcg 215
Db 4175 GGCCTCTCATCIGTACGCCGCGTGTATGTCGTCRCCCTCCAGGAGCTTACGAG 4116

Query 216 aatacagcgctgttccattcgaaaggagacaaatcgatgtttcgatca 275
Db 4115 CATAACAGATCGTCTCTTCAGTCGTCGTTGATGACCAATCACTGTCATCCGTGGR 4056

Query 276 gcttcaggatgatagaaggaaatgttacaacgtttatcgatcatcgatcat 335
Db 4055 TCCAAACAAAGGCGCAGGCAAGATCACCCGCTTACCTTAAAGATGTCCTTCCTC 3996

Query 336 atcgagagatgttaccaaggaaaaaaggaaatgttacatcc 395
Db 3995 ATGGAGCTGACCCGGAGGCCAACGGTGCCTCGCACCCCTGGTATCGACGCC 3936

Query 396 ttaatgttgc 408
Db 3935 TCTAAGGTGTC 3923

RESULT 6
SPBC29B5/c
LOCUS SPBC29B5 14147 bp DNA
DEFINITION S.pombe chromosome II cosmid c29B5.
ACCESSION AL391603
VERSION AL391603..1
KEYWORDS 60s ribosomal protein L26; atf1; isp4; sexual differentiation; transcription factor; fission yeast.

SOURCE Schizosaccharomyces pombe
ORGANISM Schizosaccharomyces pombe
Bukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

REFERENCE 1 (bases 1 to 14147)
AUTHORS Wood, V., Rajandream, M.A., Barrell, B.G., Saunders, D. and Harris, D.
TITLE Direct Submission
JOURNAL Submitted (24-AUG-1999) European Schizosaccharomyces genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk
COMMENT Notes:
Details of yeast sequencing at the Sanger Centre are available on

the World Wide Web.

(URL, http://www.sanger.ac.uk/Projects/S_pombe/)

During 1995 to 1996 about 66% of *S. pombe* chromosome 1 was sequenced by the Sanger Centre. The sequencing of the *S. pombe* genome is now being continued with funding from The European Commission. Fourteen European sequencing laboratories, including

the Sanger Centre, are participating in the project. Protein coding regions (CDS) have been predicted with the help of computer analysis using the Genefinder program in PomBase (an ACEDB database) with additional predictions for the branch-acceptor sites supplied by the program SpjSplice. CAUTION: It is possible that for any individual CDS we may have underestimated or overestimated the number of introns/exons or we may not have chosen the correct splice donor/acceptor sites.

CDS are numbered using the system eg SPBC25H2.01c. SP (S. pombe), B (chromosome 2), c25H2 (cosmid name), .01 (first CDS), c (complementary strand).

The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous. The length in codons is given for each CDS.

IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once or longer, because we arrange for a small overlap between neighbouring submissions.

Cosmid c29B5 is overlapped at the 3' end by cosmid c27, EMBL entry SPBC27, accession number AL353866.

FEATURES source

1..14147 /organism="Schizosaccharomyces pombe"
/strain="972h-"

/db_xref="taxon:4896"
/chromosome="II"
/clone="cosmid c29B5"

/map="III"
1..101 /note="SPBC29B5.01, len:567"
/codon_start=1

/label="atf1"
/product="transcription factor atf1"
/protein_id=CAC0510..1

/db_xref="GI:9955011.."
/translation="MSPPVNPNTSTEPPASA
VSGQNNQSTTATPGGASVANAPKAMP
AFTGTTMHPPLSPATYDTRPLDYNNSDASA
GAGIPAPSGTGTANLQKQPSSENQDAACT
AVGTMNPQASRTOOPMFMSQOENPMV
NGSNLNPQTSNSMPNTPMSPLKLEN
SANCJTSYKSSRNRSKNTDEERLSKLSRNQAKL
NENEILSLSAQASLREELVSLKTLIAHKDCPVA
KNSAVATSVIGSDLAQRINLG

1658..1825 /gene="isp4"
/note="Match to PFO0170 bz1P, bz1P transcription factor
Score 48 62"
complement(2997..5354)

/gene="isp4"
/note="SPBC29B5.02c"
/note="SPBC29B5..02c"
complement(2997..5354)

/gene="isp4"
/note="SPBC29B5.02c, len:786"
/codon_start=1

/label="isp4"
/product="sexual differentiation process protein isp4"
/protein_id="AAC0511..1"

/db_xref="GI:9955014.."

Query Match 26.9%; Score 109.8; DB 11; Length 979;
 Best Local Similarity 60.0%; Pred. No. 1e-18; Mismatches 183; Indels 0; Gaps 0;

Qy 303 aacaacgtgtatcgaagaatatcgatcatatcgagatgtaccagaaaaaqcg 362
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 66383 AGGGCGTGCACCGCTCAAGTGGTCATCCTCAGATAGGAGACCGCGAGAGGGC 66442
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Qy 363 aatgtaaatgtatgtttatcgatggatatactccataaacttgcggatgg 405
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 66443 AACGGTCCACCGTGGCCGTCGCATCCATCCACAGTCG 66485
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
 CNS061DT CNS061DT 979 bp DNA STS 10-JAN-2001
 DEFINITION T7 end of clone AS0AA019H06 of library AS0AA from strain CLIB 533
 ORGANISM *Saccharomyces bayanus*, sequence tagged site.

ACCESSION AL400167
 VERSION AL400167.1
 KEYWORDS STS

REFERENCE 1 (bases 1 to 979)
 AUTHORS Bon,E., Neuveline,C., Casaregola,S., Artigueneave,F., Wincker,P.,
 Aigle,M., and Durrens,P.
 TITLE Saccharomyces bayanus var. *uvvarum*
 JOURNAL FBS Lett. 487 (1), 37-41 (2000)
 PUBLISHED 1112880
 REFERENCES 2 (bases 1 to 979)
 AUTHORS Soucet,J.L., Aigle,M., Artigueneave,F., Blandin,G.,
 Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
 de-Montigny,J., Dutil,B., Durrens,P., Lepingle,A., Llorente,B.,
 Malpertuy,A., Neuveline,C., Ozier-Kalogeropoulos,O., Potier,S.,
 Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesołowski-Louvel,M.,
 Wincker,P., and Weissbach,J.
 TITLE Genomic Exploration of the Hemiascomycetous Yeasts: 1. A set of
 yeast species for molecular evolution studies(1)
 JOURNAL FBS Lett. 487 (1), 3-12 (2000)
 PUBLISHED 1112876
 REFERENCES 3 (bases 1 to 979)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (07 SEP 2000) Genoscope - Centre National de Sequencage,
 2 rue Gaston Crémieux, CP 5716, 91057 EVRY cedex, FRANCE. (E-mail :
 seqr@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 COMMENT This STS is part of a random genomic sequencing program of thirteen
 yeast species: *Saccharomyces bayanus* var. *uvvarum*, *Saccharomyces*
exiguus, *Saccharomyces servazzii*, *Zygosaccharomyces rouxii*,
Saccharomyces kluveri, *Kluveromyces thermotolerans*, *Kluveromyces*
lactis var. *lactis*, *Kluveromyces marxianus* var. *marxianus*, *Pichia*
angusta, *Debaromyces hanseni* var. *hanseni*, *Pichia sorbitophila*,
Candida tropicalis and *Yarrowia lipolytica*. Genomic inserts of 3 to
 5 kb were prepared and both extremities were sequenced. See
 keywords for description of this sequence and for the sequence of
 the other extremity of this insert.

FEATURES source
 1. .979
 /organism="Saccharomyces bayanus"
 /strain="CLIB 533"
 /variety="uvvarum"
 /db_xref="AS0AA019H06"
 /clone="AS0AA019H06"
 /clone_id="AS0AA"
 /note="end : T7"
 /note="end : T7"
 /note="end : T7"
 /note="end : T7"
 /note="Similar to *Saccharomyces cerevisiae* ORF YGR034w [
 RPL26B ; 605 large subunit ribosomal protein]"
 /evidence="not_experimental"
 misc_feature
 /note="Similar to *Saccharomyces cerevisiae* ORF YGR034w [
 RPL26B ; 605 large subunit ribosomal protein]"
 /evidence="not_experimental"
 BASE COUNT 337 a 168 c 186 g 286 t 2 others
 ORIGIN

RESULT 9
 AX073298 AX073298 1348 bp DNA PAT 25-JAN-2001
 LOCUS Sequence 409 from Patent WO0102550.
 DEFINITION Saccharomyces mitosporic Saccharomycetales; Candida.
 ACCESSION AX073298
 VERSION AX073298.1
 KEYWORDS Candida albicans.

REFERENCE 1 (bases 1 to 1448)
 AUTHORS Contreras,R.H., de Backer,M.D., Luyten,W.H., Malcorps,I.K.,
 Neijissen,B.J., and Reekmans,R.J.
 TITLE Cell death related drug targets in yeast and fungi
 JOURNAL Patent: WO 010250-A 409 11-JAN-2001;
 JANSSEN PHARMACEUTICA N.V. (BE)
 FEATURES source
 1. .1348
 /organism="Candida albicans"
 /db_xref="taxon:576"
 BASE COUNT 512 a 194 c 196 g 446 t
 ORIGIN

Query Match 26.9%; Score 109.8; DB 6; Length 1348;
 Best Local Similarity 60.0%; Pred. No. 1e-18; Mismatches 183; Conservative 0; Indels 0; Gaps 0; Gaps 0;

Qy 104 ttccaaatgtatcgatcgatcataaagaaatcgatccatccatcgaccatc 163
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 976 TTATATAGACCTTCCTTCATCTCGTCCTAACGTTAGAAAGCTTATTCCTC 1035
 Qy 164 ttctgtatcgatcgatcataatcgatcgatccatccatcgaccatc 223
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1036 AGTGGAAAAGAGTCTTATTCCTCCATTATCCAAAGAATTAGAACACATACA 1095
 Qy 224 cgtttatgtatcgatcgatccatccatcgaccatc 283
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1096 TCTCAATCTTGCCAAATAGACAAATGATGAGTTAGTTAGGTCTAAAGA 1155
 Qy 284 ggtagatcgatcgatccatccatcgaccatc 343
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1156 AGTTCGAAAGTAAAGTTATTCGTTATAGATGATGAAATTGCTATCAAGTGTATA 1215

Mon Mar 4 11:24:42 2002

us-09-540-235-1.rge